

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2002, 05:28:43 ; Search time 2658.23 Seconds

(Without alignments)  
6925.982 Million cell updates/sec

Title: US-09-602-833A-1

Perfect score: 1116  
Sequence: 1 atggagacataagtggtgtt.....ctttagccttaacttga 1116

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: 1: gb\_ba: 2: gb\_hgt: 3: gb\_in: 4: gb\_cm: 5: gb\_ov: 6: gb\_pat: 7: gb\_ph: 8: gb\_pl: 9: gb\_pr: 10: gb\_ro: 11: gb\_sts: 12: gb\_sy: 13: gb\_un: 14: gb\_vl: 15: em\_ba: 16: em\_fun: 17: em\_hum: 18: em\_in: 19: em\_om: 20: em\_or: 21: em\_ov: 22: em\_pat: 23: em\_ph: 24: em\_pl: 25: em\_ro: 26: em\_sts: 27: em\_sy: 28: em\_un: 29: em\_vl: 30: em\_hgt\_hum: 31: em\_hgt\_inv: 32: em\_hgt\_rod: 33: em\_hgt\_hum: 34: em\_hgt\_inv: 35: em\_hgt\_rod: 36: em\_hgt\_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1116	100.0	1116	6	AX063232 Sequence
2	1114.4	99.9	4860	9	AX0308569 Homo sapi
3	683.2	61.2	172966	9	AC013467 Homo sapi
4	681	61.0	681	6	AX063234 Sequence
5	211.4	18.9	145795	2	AX063234 Sequence
6	211.4	18.9	202324	2	AX063234 Sequence
7	161	14.4	202324	2	AX068720 Homo sapi
8	130.4	11.7	2036	2	AX068720 Homo sapi
9	80.2	7.2	150804	9	AC018500 Homo sapi
10	80.2	7.2	164702	33	AC016333 Homo sapi
11	80.2	7.2	166043	9	AC090947 Homo sapi
12	80.2	7.2	185608	9	AC027126 Homo sapi
13	67.2	6.0	216481	2	AC027653 Mus muscu
14	64.4	5.8	169580	2	AC092762 Homo sapi
15	63.6	5.7	1658	9	AF359380 Homo sapi
16	63.6	5.7	2156	9	BC003193 Homo sapi
17	63.6	5.7	2886	9	AF332199 Homo sapi
18	63.6	5.7	165852	2	AC092874 Sus scrofa
19	63.2	5.7	2754	2	BC003407 Homo sapi
20	63.2	5.7	2846	9	BC008586 Homo sapi
21	63.2	5.7	2847	9	AK022951 Homo sapi
22	62.8	5.6	142296	2	AP004035 Homo sapi
23	62.8	5.6	142396	9	HSAC000120 Human BAC
24	62.8	5.6	257967	2	AL365337 Mus muscu
25	62.4	5.6	133841	2	AC079378 Rattus no
26	62.4	5.6	188150	2	AC079378 Rattus no
27	58	5.2	156294	2	AC092409 Papio cyn
28	58	5.2	198746	2	AC092517 Papio cyn
29	57.4	5.1	1127	9	AY007147 Homo sapi
30	57.4	5.1	2384	6	AX099399 Sequence
31	57.4	5.1	6409	9	AF263744 Homo sapi
32	57.4	5.1	162598	2	AC079091 Homo sapi
33	57.4	5.1	174297	2	AL512367 Homo sapi
34	57.4	5.1	212127	2	AL391497 Homo sapi
35	56.8	5.1	182366	2	AC041041 Homo sapi
36	56.8	5.1	182885	2	AC087763 Homo sapi
37	56.8	5.1	182914	2	AC090567 Homo sapi
38	56.8	5.1	185510	2	AC011018 Homo sapi
39	56.2	5.0	3214	3	CEL276590 Caenorhab
40	56.2	5.0	5109	10	RNU66707 Rattus norv
41	55.8	5.0	2410	9	AK001332 Homo sapi
42	55.8	5.0	5765	9	AB033051 Homo sapi
43	55.2	4.9	3159	9	AB016816 Homo sapi
44	55.2	4.9	6125	6	AX188327 Sequence
45	54.2	4.9	5821	9	AF276423 Homo sapi

#### ALIGNMENTS

RESULT	1	LOCUS	AX063232	1116 bp	DNA	PAT	24-JAN-2001
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
FEATURES							
source							

CDS

1. 1116  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC35077.1"  
/db\_xref="GI:12541059"  
/translation="MGKRVVFDISVIRALMETHYKHKHAKMOKKEVERLEKSALEKIK  
EEMNFVACRRKGIPOAVYCKNGFIDTSVRLDKIENNTLIOTLPYIOLFOEMRIIDLPKNOISHP  
FELSGEHMTELPDLKEDTOLREMYISNTLQIIPYIOLFOEMRIIDLPKNOISHP  
AEICGLNKEKLANGYNKLKSIPELGDCEKLERDGSNTELMELPEPELKOYVE  
VDISANKNESSVPICTVLMNSIOWMIDISSNNITDIPODIDREELOSPLTKNKLTYLP  
KSMNLKKTLLVYSGDHVLELPALCDSSPRLKPYSLMNPIDNACEDGNEIMESB  
RDRHDEKEMKAYIEDLKRESEVSTTTKVSFLQ"

BASE COUNT 343 a 224 c 265 g 284 t

ORIGIN

Query Match 100.0%; Score 1116; DB 6; Length 1116;  
Best Local Similarity 100.0%; Pred. No. 9.2e-283;  
Matches 1116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggagacataagtggtgtgtcttcgcacacattctgtcatcacagagcctgttggaactcgt 60  
DB 1 ATGGAGACTAAGTGTGTGTCTTGACATTTCTGTCAATCAGAGCCTTGTGGAAACTCGT 60

QY 61 gtcaagaagcacaagcttggcagaagaagaaggttgaaaggcttgagaagaagcgcttg 120  
DB 61 GTCAAGAAGCACAAGCTTGGCAGAGAAGAGGTGGAAGAGCTTGAGAAGAGCGCCTTGG 120

QY 121 gagaagaataaaggagagatgtgaacttctgtgcccgaatgcagaagaagaaggcctcccgag 180  
DB 121 GAGAAGATAAAGGAGAGATGTGACTTTGTGGCCGAATGCAGAGAGAGGATCCCGCAG 180

QY 181 gctgtatactgcaagaatgtgtctcatatagacacagcgtgcgcttcgtgacaagaattgaa 240  
DB 181 GCTGTATACCTCAAGAATGCTTCTCATATAGACACACAGCTGCGCTTGTGACAAATGAA 240

QY 241 aggaacaacttcacaaggcagaagttcacttcccaagacagaaggaagaacgcagcgtgcg 300  
DB 241 AGGAACACTCTCACAAAGGCAAGTTCACCTCCCAAGACAGAGGCAAAACGGAGCGTGC 300

QY 301 ttgtgttgtaacttcttcgggagacgtgcagcgcgaccccgatctcatatgaagaagcag 360  
DB 301 TTTGTGTTGAACCTTTCGGGGAGCACCTGAGCGAGCTCCAGATTCATTGAGAGACGAG 360

QY 361 acacactgagagaatggttacataaagcaatccttgatitcaaatcatcctactatataat 420  
DB 361 ACACACTGAGAGAATGTTACATAAAGCAATACCTGATTCATAATCATCTCAATATATT 420

QY 421 cagttattcaagcagatgagaatctgtgactgccaataaaacccaatctcatatcttca 480  
DB 421 CAGTTATTCAAGCAGATGAGAAATCTGTGATCTGCCAAAAACCAAAATCTCAACATCTTCA 480

QY 481 gcaagaatcgtgtgtttaagaacacctgaagaactcaatgtgggttcaactatctgag 540  
DB 481 GCAGAATCGGTTGTTTGAAGAACCTGAAAGAATCAATGTGGTTTCAACATCTGAGAG 540

QY 541 agcattcctcagaatgtggagatgtgaaatctagagagacttgatgttcttggaaat 600  
DB 541 AGCAATTCCTCCAGAAATGGAGATGTGAAAATCTAGAGAGACTGATGTGTTCTGAAAT 600

QY 601 ctgaataatgaagcgtgaccttgaaatgaatgaatgaagaagcttcatctgttagat 660  
DB 601 CTGAATAATGAAGCCTGCTTGAATTAAGTAATTTGAAGCAAGTTACATTTGTAGAT 660

QY 661 atctcagaacaagaattctcagtgctccaatctgtgcttcgagatgtcgaaattgag 720  
DB 661 ATCTCAGCAAAAGATTCTCAGTGTCCCAATCTGTGCTTCGGGATGTCCGAATTTGAG 720

QY 721 tggttgatatcagcagcaataaactgacagacctgcgcgaagatatagacagcgtagag 780  
DB 721 TGGTTGATATCAGCAGCAATAAATCCTGACGACCTGCCGCAAGATATAGACAGCGCTAG 780

QY 781 gagctgcagagcttctctgtgtataaaaaaagaagtgcactcctcctatctcatgctg 840

DB 781 GAGCTGCAGAGCTTCTCTGTATAAAAACAAGTTGACCTACCTCCATTCATGCTG 840

QY 841 aacccaagaagctcactctgttactgtcaagtgaggagaccatttgtagactcccaact 900  
DB 841 AACCCAAGAAGCTCCTCTGTATCGCATGTGGGACCATTTGTGGAGTCCCAACT 900

QY 901 gcccttctgactcatccacacaccttaaatctgaaacccatbtatgaaatccatctgat 960  
DB 901 GCCCTTGTGACTCATCCACACCTTTAAATTTTGAAGCCTTATGACACATCTTATGAT 960

QY 961 aatgcccaatgtgaagatgtgcaatgaataatgtgaagaatgaaaggatcgccacatttt 1020  
DB 961 AATGCCCAATGTGAAGATGCGCAATGAATATGGAAGTGAACGGGATCGCCACATTTT 1020

QY 1021 gataaagaagtataaagccatataatgaagaccttaagaagaagaagaatctgtccagc 1080  
DB 1021 GATAAGAAGTATATAAGCCCTATATTTGAAGACCTTTAAGAAAGAAATCTGTCCAGC 1080

QY 1081 tataccaccaagtgctctttagccttcaacttga 1116  
DB 1081 TATACCACCAAGTGTCTTTAGCCTTCAACTTGA 1116

RESULT 2  
HSA308569  
LOCUS  
DEFINITION  
HSA308569.2 4860 bp mRNA PRI 06-JUN-2001  
Homo sapiens mRNA for leucine-rich repeat-containing 2 protein (LRRC2 gene).  
AJ308569  
ACCESSION  
VERSION  
KEYWORDS  
LRRC2 gene.  
leucine-rich repeat-containing 2 protein; LRRC2 gene.  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 4860)  
Kiss, H., Yang, Y., Kost-Alimova, M., Szeles, A., Kholodnyuk, I., Kedra, D., Kiss, C., Klein, G., Imreh, S., and Dumanski, J. P.  
Transcriptional map of the Common Eliminated Region 1 on human 3p21.3  
Unpublished  
2 (bases 1 to 4860)  
Kiss, H.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
SUBMITTED (27-FEB-2001) Kiss H., Microbiology and Tumorigenology Center (MTC), Karolinska Institute, Box 280, Stockholm, S-17177, SWEDEN  
REMARK  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
SUBMITTED (06-JUN-2001) Kiss H., Microbiology and Tumorigenology Center (MTC), Karolinska Institute, Box 280, Stockholm, S-17177, SWEDEN  
COMMENT  
FEATURES  
source  
On Jun 8, 2001 this sequence version replaced gi:13186111.  
Location/Qualifiers  
1. 4860  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3p21.3"  
168. 1283  
/gene="LRRC2"  
168. 1283  
/gene="LRRC2"  
/gene="LRRC2"  
/codon\_start=1  
/product="leucine-rich repeat-containing 2 protein"  
/protein\_id="CAC33442.1"  
/db\_xref="GI:13186112"  
/translation="MGKRVVFDISVIRALMETHYKHKHAKMOKKEVERLEKSALEKIK  
EEMNFVACRRKGIPOAVYCKNGFIDTSVRLDKIENNTLIOTLPYIOLFOEMRIIDLPKNOISHP  
FELSGEHMTELPDLKEDTOLREMYISNTLQIIPYIOLFOEMRIIDLPKNOISHP"

AEIGCLKNKLNVGNFYKLSIPPELGDENLERLDCSGNLELMELEPFELSNKQVTF  
 VDISANKFSSVPICTJLMDISNNLJDPDIDLEJSEFLKXNLTFLP  
 YSLNKLXLTLLVSGDLNLELPTALCDSSPTPKFVSIJMDNPIJDNQCEJGNIMSE  
 RDRQHDKEKMYIEDLKERESVPSYTTKVSISLOL"

BASE COUNT 1480 a 924 c 1162 g 1294 t  
 ORIGIN

Query Match 99.9% Score 1114.4: DB 9: Length 4860:  
 Best Local Similarity 99.9%: Pred. No. 3.1e-282;  
 Matches 1115: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggagacataaagtgtgtgtcttcgacattctgtlcatcagagcctgtgtgaaactcgt 60  
 DB 168 ATGGAGACATAAAGTGTGTCTTCTGACATTCTCTCATCAGAGCCTGTGGGAACGCT 227  
 QY 61 gtcaagaagcacaagcttggcagagaagaagggtggaaggcttgaagagcgcttg 120  
 DB 228 GTCAAGAGACCAAGCTTGGCAGAGAGAGAGGCTGGAAGGCTTGAAGAGAGCGCTTG 287  
 QY 121 gagaagaataaggagaggtgtggaacttggccgaatgcagaggaaggagcattccccag 180  
 DB 288 GAGAGATTAAGAGAGAGGTGGAATGTCGCCGAATGCAAGAGAGAGGAGGATCCCGCAG 347  
 QY 181 gctgtactgcagaagatggtctcatagacacagcgctgcgctcttgcgaagaattgaa 240  
 DB 348 GCTGTATCTGCAAGATGGCTTCATGACACCGCTGCGCTTCTTGACAGATTTGAA 407  
 QY 241 aggaacacatcacaaggcagagttcaatcccaagagacagaggaacagagagtcg 300  
 DB 408 AGGAACACTCTCAACAGCAGAGTTCCTCCCAAGAGACAGAGCAACGAGCAGTGGCG 467  
 QY 301 tttgtgttgaactttcttggggagcactgtgacgagctcccaagattatgaagagcag 360  
 DB 468 TTTGTGTTTAACCTTTCTGGGAGCACTGAGCGAGCTCCAGATTCATTTGAAGAGCAG 527  
 QY 361 acacacactgagagaatgtlcatagaacaatacctgtatcaaatcattccatacatatt 420  
 DB 528 ACACACTGAGAGATGTCATAGCAATACCTGATTCATTAATCTCTACATATATT 587  
 QY 421 cagttattcaagcagatgagaaattctgagatcgcacaaacaaatccaacatccaca 480  
 DB 588 CAGTTATTTCAAGAGATGAAATTCGATCTGCCAAAAAACCAATCTCATCACTTCCA 647  
 QY 481 gcggaatcggtgttgaagaacctgaaagaaactcaatgtggtttcaactatcgaag 540  
 DB 648 GCGAATAATCGGTGTGGAAGAACCTGAAGAACTCATGTGGCTTCAACTATCTGAAG 707  
 QY 541 agacttcccaagaattgagagattgtgaaatctagaagacatgagatcttctcgaat 600  
 DB 708 AGCAATCTCCAGATTTGGAGATTGTGAATAATCTAGAGAGCTGGATTTGTCGAAAT 767  
 QY 601 ctagaatlaatgagctgaccttgaattgaatattgaagcaagttacattttagat 660  
 DB 768 CTAGAATTATATGAGAGCTGCCCTTTGAATTAATTTGAAGCAAGTTACATTGTGAT 827  
 QY 661 atcttcagcaacaagtttccagtgctccaatctgttctctggagatgctgaattcag 720  
 DB 828 ATCTCAGCAACAAGTTTCCAGTGCTCCAAATCTGTGTCTGCGAGTGTGAATTTTCAG 887  
 QY 721 tggttgataatcagcaacaatacctgaacagcctgcgcaagaatataagaagctagag 780  
 DB 888 TGGTGATATTCAGACACATTAACCTGACGACCTGCCGAGATATAGACAGCTTAGAG 947  
 QY 781 gagctcagagcttctctgtataaaacaagttgacctaccctaccctatccatgctg 840  
 DB 948 GAGCTGAGAGCTTCTCTGTATAAAAACAAGTTGACCTTCCCTTATTCATCTG 1007  
 QY 841 aaactaagaagctcactcgttagtgcagagggaggaacatttggggagctcccaact 900  
 DB 1008 AACCTGAGAGAGCTCACTCTGTTAGTGTGAGGAGCAATTTGGTGAGAGCTCCCACT 1067  
 QY 901 gcccttgtgactcaccacattaaattgtlaagccttataagccttataagcctatgat 960

DB 1068 GCCCTTGTGACTCATCCACACCTTTAAATTTGAAGCCTTATGAGCAATCTATGAT 1127  
 QY 961 aatgcccaatgtgaagaatgcaatgaataatgaaagttgaagagagctgcacattt 1020  
 DB 1128 AATGCCCAATGTGAAGATGGCAATGAATAATGGAAGTGAACGGAGATCCGCAACATTTT 1187  
 QY 1021 gataagaagttatgaagaagcctatattgaagaccttaagaagaagaatcgttccagc 1080  
 DB 1188 GATTAAGAGATTATGAAGCCTTATTTGAGACCTTTAAAGAAAGATCTGTTCCAGC 1247  
 QY 1081 tataccaacaagtgcttctttagcttccatcttga 1116  
 DB 1248 TATACCAACCAAGTGTCTTTAGCTTCACACTTTGA 1283

RESULT 3  
 AC013467/c  
 LOCUS AC013467 172966 bp DNA PRI 25-MAY-2001  
 DEFINITION Homo sapiens clone RP11-451F14, complete sequence.  
 ACCESSION AC013467  
 VERSION AC013467.8 GI:14196420  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Waterston,R.H.  
 TITLE 1 (bases 1 to 172966)  
 JOURNAL The sequence of Homo sapiens clone

REFERENCE  
 AUTHORS Waterston,R.H.  
 TITLE 2 (bases 1 to 172966)  
 JOURNAL Direct Submission  
 Submitted (12-NOV-1999) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA

REFERENCE  
 AUTHORS Waterston,R.H.  
 TITLE 3 (bases 1 to 172966)  
 JOURNAL Direct Submission  
 Submitted (25-MAY-2001) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA

COMMENT  
 Center project name: H\_NH0451F14.  
 On May 25, 2001 this sequence version replaced gi:13624417.

FEATURES  
 source  
 Location/Qualifiers  
 1..172966  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="RP11-451F14"

BASE COUNT 57371 a 34327 c 31782 g 49486 t  
 ORIGIN

Query Match 61.2%: Score 683.2: DB 9: Length 172966;  
 Best Local Similarity 86.2%: Pred. No. 1.4e-168;  
 Matches 828: Conservative 0; Mismatches 118; Indels 14; Gaps 6;

QY 1 atggagacataaagtgtgtgtcttcgacattctgtlcatcagagcctgtgtgaaactcgt 60  
 DB 169367 ATGGAGATACAAAGTGCATATTGACATTTCTGTCTCAGAGCCTTGTGGGAACCTGT 169308  
 QY 61 gtcaagaagcacaagcttggcagagaagaagggtggaaggcttgaagagcgcttg 120  
 DB 169307 GTCAAGAGACAC- AAGTTTGGCAGAAAGAGGCAAGAAAGTCTTGAAGAGTGCATCA 169249  
 QY 121 gagaagaataaggagaggtgtggaacttggccgaatgcagaggaaggagcattccccag 180  
 DB 169248 GAAAAGCAAGAGAGAGAGTGGCCCTTGGAGGTCAAGTGCAGAGAGAAAGCATCTCTCG 169189  
 QY 181 gctgtata---ctgcaagaatggtctcatagaca---caagcgtgcgcttctgacaa 233  
 DB 169188 GCTGAGTACTGCGTGGGATAGTGGCTTCATATGACATTCATAGCATGCAGCTTTCGAAAA 169129

QY	234	gattgaaaggaaacattccacaaggcagaagtttactctcccaaggccagaggaacagag	293
Db	169128	GATAGAAAGGAAAGGCTCTCACAAAGGCAAGATTCACCTTCCCAAGGACAGAGGCAAAATGTGAG	169069
QY	294	cagtcgcttctgtttgtaactcttcctggggaagcagcgagaccccaagcttcaatga	353
Db	169068	CAGTGTGTTTGTGTGGACTTTCTCGGGGAGTACTGAGAGAGCTGCCAGATTTGTAA	169009
QY	354	ggagcagacacacaccttgaagaaatggatataaagaacataacctgtatccaatcattcc	413
Db	169008	GGAGCAGACGCACTCTGAGAGAAATGGCAGTATGCGACACCTCGATTCAAAATTA	168949
QY	414	atatactcaagtattctcaagcgaatgaatcttgaatcttgcacaaaaaccaaattcc	473
Db	168948	ATATATTGAGCTGTTTCAAGCATGATGAAATTCAGATCTGCCAAAGAAACAAATCTCA	168889
QY	474	ctcttcacagacaagaatcagttgtcttgaagaaccttgaagaacccaatctgtggctca	533
Db	168888	TTTTCCACTTAACCTCACTGTTGTTTGAAGACCTGGAATTAACCTATGTGAGTTTCA	168829
QY	534	ctctgaagacattccctccagaatctggagaaatgtgaaatactagaagacatgattgtc	593
Db	168838	TCTGAAAGACATCTCTCCAGACCTGGAGATGTG-AAATCTAAGAGGCACTGATTTG	168770
QY	594	tggaaattcagaatataaggaagctgcctcttgaattgaatattgaagcaagttac	653
Db	168769	TGGAAATCTGGAATTCACAGACCTCCCTTTGAATTAAGTAATGGAAGCCAGTTGCAT	168710
QY	654	tgatagatctccaagaacaagtttccagttgctccaaatctgtgctcgcagatctg	713
Db	168709	CGTACATATCTCAGCAAAACAACTTTTCCAGTGTCCCAAGTCTGTGCTGAGGATGTC	168651
QY	714	tttcagctggttgaatcagcaacaataacctgacccgacccgtccgcaagatataag	773
Db	168650	TTTGCAGTGTGTGATATCAGCAACAAATTAACCTGACTGCTTGCAGCAAGAT----	168595
QY	774	gctagaagagctgcagagcttctcttgtataaacaagttgacctcttccattc	833
Db	168594	GCTGGAAGAGCGCAGGGGCTTTCTTTGATAGAAGCAAGTTGACCTTACCTTACCC	168535
QY	834	catgttgaaacctgaagaagctcacctctgtatgctcagttgggaacattgttgaagct	893
Db	168534	CATGCTTAACCTGGAAGAGCGACCTTGTATTCGTCAATGGGACACACTTGTGTGAGCT	168475
QY	894	cccaactgcctcttgtgactcaaccacacccctttaaattgttaagccctatgacaatcc	953
Db	168474	TTCACTGTCTGTGTGTGACTCTCCACACACTTAAATTTGTAAGCTTATAACATCC	168415
RESULT	4		
AX063234	AX063234	681 bp	DNA
LOCUS	Sequence	3 from Patent	W00078959.
DEFINITION	AX063234		
ACCESSION	AX063234.1	GI:12541060	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
	human.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE			
ADTHORS	1 (bases 1 to 681)		
TITLE	Tuner,A.C., Zambrowicz,B., Nehls,M., Friedrich,G.A. and Sands,A.T		
JOURNAL	Human genes and proteins encoded thereby		
	Patent: WO 0078959-A 3 28-DEC-2000;		
	Lexicon Genetics Incorporated (US)		
FEATURES			
source			
	1..681		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	1..681		
	/note="unamed protein product"		
	/codon_start=1		

[illegible]

ORGANISM Homo sapiens  
 Eukaryota: Metazoa: Chordata: Craniata: Vertebrata; Euteleostomi;  
 Mammalia: Eutheria: Primates: Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 145795)  
 AUTHORS Kiss, H., Yang, Y., Kost-Alimova, M., Szekes, A., Kholodnyuk, I.,  
 Kedra, D., Kiss, C., Klein, G., Imreh, S., and Dumnanski, J.P.  
 TITLE Transcriptional map of the common eliminated region 1 on human  
 3p21.3  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 145795)  
 AUTHORS Kiss, H.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-APR-2001) Kiss H., Microbiology and Tumorbiology  
 Center (MTC), Karolinska Institute, Box 280, Stockholm, S-17177,  
 SWEDEN  
 COMMENT The sequencing contigs are in order and the gaps between them are  
 represented by 100 n's.  
 Contig 1: 1-2508 Contig 2: 2609-17434 Contig 3: 17535-21987  
 Contig 4: 22088-23066 Contig 5: 23167-28845  
 Contig 6: 28946-30840 Contig 7: 30941-37879 Contig 8:  
 37980-38928 Contig 9: 39029-40834 Contig 10: 40935-44527 Contig  
 11: 44628-48077 Contig 12: 48178-52755 Contig 13: 52856-53293  
 Contig 14: 53394-54768 Contig 15: 54869-57895 Contig 16:  
 57996-59322 Contig 17: 59423-60159 Contig 18: 60260-61231  
 Contig 19: 61332-62928 Contig 20: 63029-63852 Contig 21:  
 63953-75555 Contig 22: 75656-77760 Contig 23: 77861-79686  
 Contig 24: 79787-80857 Contig 25: 80958-87622 Contig 26:  
 87723-89164 Contig 27: 89265-90007 Contig 28: 90108-92035  
 Contig 29: 92136-97298 Contig 30: 97399-102032 Contig 31:  
 102133-106605 Contig 32: 106706-108263 Contig 33: 108364-110022  
 Contig 34: 110123-113747 Contig 35: 113848-125151 Contig 36:  
 125252-126026 Contig 37: 126127-129649 Contig 38: 129750-132160  
 Contig 39: 132261-133125 Contig 40: 133226-135661 Contig 41:  
 135762-142148 Contig 42: 142249-145795.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 42 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* been provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 2508: contig of 2508 bp in length  
 \* 2509 2608: gap of 100 bp  
 \* 2609 17434: contig of 14826 bp in length  
 \* 17435 17534: gap of 100 bp  
 \* 17535 21987: contig of 4453 bp in length  
 \* 21988 22087: gap of 100 bp  
 \* 22088 23066: contig of 979 bp in length  
 \* 23067 23166: gap of 100 bp  
 \* 23167 28845: contig of 5679 bp in length  
 \* 28846 28945: gap of 100 bp  
 \* 28946 30840: contig of 1895 bp in length  
 \* 30841 30940: gap of 100 bp  
 \* 30941 37879: contig of 6939 bp in length  
 \* 37880 37979: gap of 100 bp  
 \* 37980 38928: contig of 949 bp in length  
 \* 38929 39028: gap of 100 bp  
 \* 39029 40834: contig of 1806 bp in length  
 \* 40835 40934: gap of 100 bp  
 \* 40935 44527: contig of 3593 bp in length  
 \* 44528 44627: gap of 100 bp  
 \* 44628 48077: contig of 3450 bp in length  
 \* 48078 48177: gap of 100 bp  
 \* 48178 52755: contig of 4578 bp in length  
 \* 52756 52855: gap of 100 bp  
 \* 52856 53293: contig of 438 bp in length  
 \* 53294 53393: gap of 100 bp  
 \* 53394 54768: contig of 1375 bp in length  
 \* 54769 54868: gap of 100 bp  
 \* 54869 57895: contig of 3027 bp in length  
 \* 57896 57995: gap of 100 bp

```

* 57996 59322: contig of 1327 bp in length
* 59323 59422: gap of 100 bp
* 59423 60159: contig of 737 bp in length
* 60160 60259: gap of 100 bp
* 60260 61231: contig of 972 bp in length
* 61232 61331: gap of 100 bp
* 61332 62928: contig of 1597 bp in length
* 62929 63028: gap of 100 bp
* 63029 63852: contig of 824 bp in length
* 63853 63952: gap of 100 bp
* 63953 75555: contig of 11603 bp in length
* 75556 75655: gap of 100 bp
* 75656 77760: contig of 2105 bp in length
* 77761 77860: gap of 100 bp
* 77861 79686: contig of 1826 bp in length
* 79687 79786: gap of 100 bp
* 79787 80857: contig of 1071 bp in length
* 80858 80957: gap of 100 bp
* 80958 87622: contig of 6665 bp in length
* 87623 87722: gap of 100 bp
* 87723 89164: contig of 1442 bp in length
* 89165 89264: gap of 100 bp
* 89265 90007: contig of 743 bp in length
* 90008 90107: gap of 100 bp
* 90108 92035: contig of 1928 bp in length
* 92036 92135: gap of 100 bp
* 92136 97298: contig of 5163 bp in length
* 97299 97398: gap of 100 bp
* 97399 102032: contig of 4634 bp in length
* 102033 102132: gap of 100 bp
* 102133 106605: contig of 4473 bp in length
* 106606 106705: gap of 100 bp
* 106706 108263: contig of 1558 bp in length
* 108264 108363: gap of 100 bp
* 108364 110022: contig of 1659 bp in length
* 110023 110122: gap of 100 bp
* 110123 113747: contig of 3625 bp in length
* 113748 113847: gap of 100 bp
* 113848 125151: contig of 11304 bp in length
* 125152 125251: gap of 100 bp
* 125252 126026: contig of 775 bp in length
* 126027 126126: gap of 100 bp
* 126127 129649: contig of 3523 bp in length
* 129650 129749: gap of 100 bp
* 129750 132160: contig of 2411 bp in length
* 132161 132260: gap of 100 bp
* 132261 133125: contig of 865 bp in length
* 133126 133225: gap of 100 bp
* 133226 135661: contig of 2436 bp in length
* 135662 135761: gap of 100 bp
* 135762 142148: contig of 6387 bp in length
* 142149 142248: gap of 100 bp
* 142249 145795: contig of 3547 bp in length.
FEATURES
  source
    1..145795
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="3"
    /map="3p21.3"
    /clone="RP6-91P17"
BASE COUNT 39764 a 33207 c 31499 g 37225 t 4100 others
ORIGIN
Query Match 18.9%; Score 211.4; DB 2; Length 145795;
Best Local Similarity 95.2%; Pred. No. 1.7e-44;
Matches 218; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 115 gcccttgagaagaataaaggagagtggaactttgtgcccgaatgcagaaggagcgcac 174
Db 73663 GCCCTGTTTATGAAGGAGAGTGAACCTTTGTGCGCCGATCCAGAGAGAGGCATC 73604
QY 175 ccccaagctgatactgaagaatgagctcatagacaccagcgtgcgcttcgtgacaag 234

```

|||||  
Db 73603 CCCGAGGCTGTAATGCAAGATGCTTCATAGACACAGGCTCGCTTGTGACAG 73544  
QY 235 attgaagaacactctcacagagatcactcccaagacagagcaacagagc 294  
|||||  
Db 73543 ATTGAAGACACTCTCACAAGGAGATTACTTCCCAAGACAGAGCAAGGAGAC 73484  
QY 295 agtcgcttctgttgtaactctctgggagacagcagagcctccag 343  
|||||  
Db 73483 AGTCGCTTGTGTTGTAACCTTCTGGGAGCACTGAGCGTGATCGAG 73435  
|||||  
RESULT 6  
AC068720/c 202324 bp DNA HTG 07-JUL-2000  
LOCUS Homo sapiens chromosome 3 clone RP11-509121, WORKING DRAFT  
DEFINITION  
SEQUENCE, 20 unordered pieces.  
AC068720  
AC068720.2 GI:8469022  
VERSION HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 202324)  
AUTHORS Waterston, R.H.  
JOURNAL The sequence of Homo sapiens clone  
REFERENCE 2 (bases 1 to 202324)  
AUTHORS Waterston, R.H.  
JOURNAL Direct Submission  
Submitted (07-MAY-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
COMMENT On Jun 12, 2000 this sequence version replaced gi:7715661.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
----- Project Information -----  
Center project name: H\_NH0509121  
----- Summary Statistics -----  
Sequencing vector: M13; 100%  
Sequencing vector: plasmid; 0%  
Chemistry: Dye-Primer ET; 100% of reads  
Chemistry: Dye-terminator Big Dye; 0% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 193150 bases at least Q40  
Consensus quality: 195686 bases at least Q30  
Consensus quality: 197015 bases at least Q20  
Insert size: 200000; agarose-fp  
Insert size: 200424; sum-of-contigs  
Quality coverage: 4.67 in Q20 bases; agarose-fp  
Quality coverage: 4.72 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 20 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1576: contig of 1576 bp in length  
\* 1577 1676: gap of unknown length  
\* 1677 3263: contig of 1587 bp in length  
\* 3264 3363: gap of unknown length  
\* 3364 5443: contig of 2080 bp in length  
\* 5444 5543: gap of unknown length  
\* 5544 8196: contig of 2653 bp in length  
\* 8197 8297: gap of unknown length  
\* 8297 12038: contig of 3742 bp in length

\* 12039 12138: gap of unknown length  
\* 12139 15992: contig of 3654 bp in length  
\* 15993 16092: gap of unknown length  
\* 16093 21192: contig of 5100 bp in length  
\* 21193 21292: gap of unknown length  
\* 21293 24767: contig of 3474 bp in length  
\* 24767 24867: gap of unknown length  
\* 24867 30824: contig of 5958 bp in length  
\* 30825 30925: gap of unknown length  
\* 30925 37153: contig of 6229 bp in length  
\* 37154 37253: gap of unknown length  
\* 37254 45263: contig of 8009 bp in length  
\* 45263 45363: gap of unknown length  
\* 45363 56733: contig of 11371 bp in length  
\* 56734 56833: gap of unknown length  
\* 56834 72157: contig of 15323 bp in length  
\* 72157 72257: gap of unknown length  
\* 72257 84702: contig of 12446 bp in length  
\* 84703 84802: gap of unknown length  
\* 84803 101052: contig of 16250 bp in length  
\* 101053 101152: gap of unknown length  
\* 101153 117709: contig of 16557 bp in length  
\* 117710 117809: gap of unknown length  
\* 117810 132650: contig of 14841 bp in length  
\* 132651 132750: gap of unknown length  
\* 132751 152071: contig of 19321 bp in length  
\* 152072 152171: gap of unknown length  
\* 152172 176190: contig of 24019 bp in length  
\* 176191 176291: gap of unknown length  
\* 176291 202324: contig of 26034 bp in length.

FEATURES  
Source  
1..202324  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-509121"  
1..1576  
/note="assembly\_name:Contig4"  
1677..3263  
/note="assembly\_name:Contig5"  
3364..5443  
/note="assembly\_name:Contig6"  
5544..8196  
/note="assembly\_name:Contig7"  
8297..12038  
/note="assembly\_name:Contig8"  
12139..15992  
/note="assembly\_name:Contig9  
clone\_end:17  
vector\_side:right"  
16093..21192  
/note="assembly\_name:Contig10"  
21293..24766  
/note="assembly\_name:Contig11"  
24867..30824  
/note="assembly\_name:Contig12"  
30925..37153  
/note="assembly\_name:Contig13"  
37254..45262  
/note="assembly\_name:Contig14"  
45363..56733  
/note="assembly\_name:Contig15"  
56834..72156  
/note="assembly\_name:Contig16"  
72257..84702  
/note="assembly\_name:Contig17"  
84803..101052  
/note="assembly\_name:Contig18"  
101153..117709  
/note="assembly\_name:Contig19"  
117810..132650  
/note="assembly\_name:Contig20  
clone\_end:SP6

```

misc_feature      vector_side:left"
                  132751..152071
                  /note="assembly_name:Contig21"
misc_feature      152172..176190
                  /note="assembly_name:Contig22"
misc_feature      176291..202324
                  /note="assembly_name:Contig23"
BASE COUNT      56787 a 45300 c 43765 g 54528 t 1944 others
ORIGIN
Query Match      18.9%; Score 211.4; DB 2; Length 202324;
Best Local Similarity 95.2%; Pired. No. 1.8e-44;
Matches 218; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 115 gccctggagaagaataaaggagagtgtaacttggccgaatgcaagaagaagcattc 174
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 30432 GCCCTGTTTAAAGTAAAGGAGAGTGAACCTTGTGGCCGAAATGACAGAGGCATC 30373
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 175 ccccaagctgtatactcaagaatggtcatagacacacagcgtgctgtcttgacaag 234
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 30372 CCCAGGCTGTATCTGCAAGATGCTTCATAGACACACGCGGCTTCTTGACACAG 30313
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 235 attgaaagacacactcacaagcagagtcacttcccaagagcagaagcacaagcagc 294
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 30312 ATTGAAGAGCACTCTCACAAGCAGAGTTCACTCCCAAGGACAGAGCAACGAGAGC 30253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 295 agtgcgttgtgttgtaacttctgggagacactggagcagagctccag 343
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 30252 AGTGCCTTGTGTGAACCTTCTGGGAGACACTGAGCGTGAAGTGGAG 30204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
AC068720 AC068720 202324 bp DNA HTG 07-JUL-2000
LOCUS      Homo sapiens chromosome 3 clone RP11-509121, WORKING DRAFT
DEFINITION
SEQUENCE, 20 unordered pieces.
AC068720 AC068720 GI:8469022
VERSION    HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS   human.
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 202324)
            Waterston,R.H.
REFERENCE   2 (bases 1 to 202324)
            The sequence of Homo sapiens clone
            Unpublished
            Waterston,R.H.
            Direct Submision
            Submitted (07-MAY-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            On Jun 12, 2000 this sequence version replaced gi:7715661.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0509121
----- Summary Statistics -----
Sequencing vector: p13: 100%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 193150 bases at least Q40
Consensus quality: 195686 bases at least Q30
Consensus quality: 197015 bases at least Q20
Insert size: 200000; agarose-IP
Insert size: 200424; sum-of-ctrls

```

```

Quality coverage: 4.67 in Q20 bases; agarose-IP
Quality coverage: 4.72 in Q20 bases; sum-of-ctrls
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1576: contig of 1576 bp in length
1577 1676: gap of unknown length
1677 3263: contig of 1587 bp in length
3264 3363: gap of unknown length
3364 5443: contig of 2080 bp in length
5444 5543: gap of unknown length
5544 8197: contig of 2653 bp in length
8197 8297: gap of unknown length
8297 12038: contig of 3742 bp in length
12039 12139: gap of unknown length
12139 15992: contig of 3854 bp in length
15993 16092: gap of unknown length
16093 21192: contig of 5100 bp in length
21193 21292: gap of unknown length
21293 24766: contig of 3474 bp in length
24767 24866: gap of unknown length
24867 30824: contig of 5958 bp in length
30825 30924: gap of unknown length
30925 37153: contig of 6229 bp in length
37154 37254: gap of unknown length
37254 45262: contig of 8009 bp in length
45263 45363: gap of unknown length
45363 56733: contig of 11371 bp in length
56734 56833: gap of unknown length
56834 72156: contig of 15333 bp in length
72157 72256: gap of unknown length
72257 84702: contig of 12446 bp in length
84703 84802: gap of unknown length
84803 101052: contig of 16250 bp in length
101053 101152: gap of unknown length
101153 117709: contig of 16557 bp in length
117710 117810: gap of unknown length
117810 132650: contig of 14841 bp in length
132651 132750: gap of unknown length
132751 152071: contig of 19321 bp in length
152072 152171: gap of unknown length
152172 176190: contig of 24019 bp in length
176191 202324: gap of unknown length
176291 202324: contig of 26034 bp in length.

FEATURES
source
1..202324
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="3"
    /clone="RP11-509121"
1..1576
    /note="assembly_name:Contig4"
1677..3263
    /note="assembly_name:Contig5"
3364..5443
    /note="assembly_name:Contig6"
5544..8196
    /note="assembly_name:Contig7"
8297..12038
    /note="assembly_name:Contig8"
12139..15992
    /note="assembly_name:Contig9
vector_side:right"
16093..21192
    /note="assembly_name:Contig10"
21293..24766

```

```
misc_feature /note="assembly_name:Contig11"
24867..30824
/note="assembly_name:Contig12"
30925..37153
/note="assembly_name:Contig13"
37254..45262
/note="assembly_name:Contig14"
45363..56733
/note="assembly_name:Contig15"
56834..72156
/note="assembly_name:Contig16"
72257..84702
/note="assembly_name:Contig17"
84803..101052
/note="assembly_name:Contig18"
101153..117709
/note="assembly_name:Contig19"
117810..132650
/note="assembly_name:Contig20
clone_end:SP6
vector_side:left"
misc_feature /note="assembly_name:Contig21"
132751..152071
/misc_feature /note="assembly_name:Contig22"
152172..176190
/misc_feature /note="assembly_name:Contig23"
176291..202324
/note="assembly_name:Contig23"
BASE COUNT 56787 a 45300 c 43765 g 54528 t 1944 others
ORIGIN
Query Match 14.4% Score 161; DB 2; Length 202324;
Best Local Similarity 97.0% Pred. No. 3.2e-31;
Matches 164; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 772 aggtcagagagagcgcagagctctctgtatataaacaagttgactactcccat 831
|||||
Db 96876 AGGCTAGAGAGAGCTGCAGAGCTTCTCTGTATATAAACAAGTTGACCTTCCCTAT 96935
|||||

QY 832 tccatgctgaacctgaagaagctcactctgttagctgagtgaggaccattgtgag 891
|||||
Db 96936 TCCATGCTGAACCTGAAGAAGCTCAGCTGTAGTCGTCAGTGGGAGACATTGGTGAG 96995
|||||

QY 892 ctcccaactgcctcttctgtgactatccacaccttaaatgttaagcc 940
|||||
Db 96996 CTCCCAACTGCCCTTGTGACTCATCCACCTTTAAAGTAGTAGCC 97044
|||||

RESULT 8
AC021919 2056 bp mRNA PRI 29-SEP-2000
LOCUS AK021919
DEFINITION Homo sapiens cDNA FLJ11857 fls, clone HEMBA1006807, moderately
similar to Homo sapiens mRNA for SPO.
ACCESSION AK021919
VERSION AK021919.1 GI:10433216
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens embryo, 10 weeks whole embryo, mainly head cDNA to
mRNA, clone_id:HEMBA1 clone:HEMBA1006807.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
Tanase,T., Nomura,Y., Togiyasu,S., Komai,F., Hara,R., Takeuchi,K.,
Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Makanatsu,A., Nakamura,Y., Nagahari,K., Masuno,Y. and Oshima,A.
NEDO human cDNA sequencing project
TITLE UNPUBLISHED (2000)
JOURNAL 2 (bases 1 to 2056)
REFERENCE Isogai,T. and Otsuki,T.
AUTHORS Direct Submission
JOURNAL Submitted (23-AUG-2000) to the DDBJ/EMBL/Genbank databases. Takao
```

```
COMMENT Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:genom@icse.hri.co.jp,
Tel:81-438-52-3391, Fax:81-438-52-3352)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source Location/Qualifiers
1..2056
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEMBA1006807"
/clone_id="HEMBA1"
/dev_stage="embryo, 10 weeks"
/tissue_type="whole embryo, mainly head"
/note="cloning vector: pME185FL3"
247..1422
/note="unnamed protein product"
/codon_start=1
/protein_id="BAJ3937.1"
/db_xref="GI:10433217"
/translation="MSREPTPLPQDKMSTGPIAESKCYOVKKYSYMTINNFSRC
REMEGEVLKSTSTSGPSDKMKWCLRNPGDLDESDYSLVLLVSCRSKSEVRAF
KFSLNKRRETRAMESQRAYRREVQDKWCFKFTIRDFLLDEANGLLPDKLTFCF
VSVVDVSNISGHTNTWTLEKPECLAEEDGKLNMENTRFDFCSFAGFGFKAHKSVL
AARSPVNAMEHEHMERESKKRVEINDLDEVEKEMRFYTGRAPLDKMAINLAA
ADRYALERLKYMCBEALCSNLSEVNAVDTVLADIHSABOLKQAIIDFIRCSVLRD
GCKDGRKWNNSQATDIESRDRDHPFKEVMAKVIEDLKERESVPSTTTVSVSLQI"
BASE COUNT 642 a 394 c 495 g 525 t
ORIGIN
Query Match 11.7% Score 130.4; DB 9; Length 2056;
Best Local Similarity 99.2% Pred. No. 1.6e-23;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 985 gaatatggaagaatggaagcgatcgccaacatttgtataaagaagtatatgaagccat 1044
|||||
Db 1291 GACATTAATGGAAGATGAAGCGGATCGCCACATTTGTGATAAGAAGTTAGAAAGCGCAT 1350
|||||

QY 1045 attgaagaccttaagaagaagaatcgtgtccagctataccaccaagtgctttagc 1104
|||||
Db 1351 ATTGAAGACCTTTAAAGAAGAATCTGTCCAGCTATACACCAAGAAGTCTTTTAGC 1410
|||||

QY 1105 cttaaccttga 1116
|||||
Db 1411 CTTCACTTTGA 1422
|||||

RESULT 9
AC018500/c 150804 bp DNA PRI 19-MAR-2001
LOCUS AC018500/c
DEFINITION Homo sapiens chromosome 3 clone RP11-275J11 map 3p, complete
sequence.
ACCESSION AC018500
VERSION AC018500.3 GI:13378163
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 150804)
AUTHORS Wu,Q., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J.,
Ding,H., Dong,W., Fan,H., Huang,X., Gong,J., Guan,Q., Gu,X., Guo,D.,
Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C.,
Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B.,
Liu,Y., Li,W., Li,W., Li,Y., Luo,C., Luo,J., Niu,Y., Qi,Q., Qi,X.,
Song,L., Song,S., Sun,M., Sun,W., Sun,X., Tan,X., Tao,R., Wang,H.,
Wang,J., Wang,J., Wang,L., Wang,L., Wang,L., Wang,R., Wang,X.,
Wang,X., Wang,Y., Wu,D., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X.,
```



YU, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.  
 Chromosome 3p genomic sequence  
 Unpublished  
 2 (bases 1 to 150804)

He, L., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, C., Luo, J., Niu, Y., Qi, Q., Qi, L., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Direct Submission  
 Submitted (13-DEC-1999) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China

REFERENCE  
 AUTHORS

3 (bases 1 to 150804)  
 Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, C., Luo, J., Niu, Y., Qi, Q., Qi, L., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Direct Submission  
 Submitted (19-MAR-2001) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China  
 On Mar 19, 2001 this sequence version replaced g1:6684196.

COMMENT

Center: Beijing Center  
 Center code: Beijing  
 Website: http://hgsc.igtp.ac.cn  
 http://www.genomics.org.cn  
 Contact: hgsc@igtp.ac.cn

Project Information  
 Center project name: RP11-275J11  
 Center clone name: RP11-275J11

Summary Statistics  
 Sequencing vector: pUC18; 100% of reads  
 Chemistry: Dye-terminator; ET 55% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 152846 bases at least Q40  
 Consensus quality: 153558 bases at least Q30  
 Consensus quality: 153695 bases at least Q20  
 Insert size: 150804; sum-of-contigs  
 Quality coverage: 10.96x in Q20 bases; sum-of-contigs

FEATURES

Location/Qualifiers  
 1. 150804  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /map="3p"  
 /clone="RP11-275J11"

BASE COUNT  
 ORIGIN 40688 a 33310 c 33710 g 43096 t

Query Match 7.2%; Score 80.2; DB 9; Length 150804;  
 Best Local Similarity 73.0%; Pred. No. 5.4e-10;  
 Matches 103; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 137 aggggaacttggccgagatgcagggagggagcattccccagcgtctatctgcaaga 196  
 Db 28318 AATTAACCTATTTTAAAAAATTCAGAGTAAGACATTCCTCCAGCCATGCTCCGCAAGA 28259  
 QY 197 atggttcataagacaccagcgtgcggtctctgcaagaattgaagaagaacctctcaaa 256

Db 28258 ATGACTTCACAGATACACGCGCTTGAAGATGAGAAAGATGAGAAACCTCAGCGAA 28199  
 QY 257 ggcagagtcacctcccaagg 277  
 Db 28198 GGTAGAGCTCAGCTTCCACAG 28178

RESULT 10

AC016333/C standard; DNA; HTG; 164702 BP.

AC016333;

AC016333.5

26-NOV-1999 (Rel. 61, Created)  
 26-MAY-2000 (Rel. 63, Last updated, Version 7)

Homo sapiens clone RP11-30G23, WORKING DRAFT SEQUENCE, 17 unordered pieces.

HTG; HTGS-DRAFT; HTGS-PHASE1.

Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

[1]  
 1-164702  
 RA Birren B., Linton L., Nusbaum C., Lander E.;  
 RT "Homo sapiens chromosome, clone RP11-30G23",  
 RL Unpublished.

[2]  
 1-164702  
 RA Birren B., Linton L., Nusbaum C., Lander E., Allen N., Anderson M.,  
 RA Baldwin J., Barna N., Beckerly R., Boguslavsky L., Boukhalter B.,  
 RA Brown A., Castle A., Colangelo M., Collins S., Collymore A., Cooke P.,  
 RA Dearlano K., Dewar K., Domino M., Donegan L., Doyle M., Ferreira P.,  
 RA Fitzhugh W., Forrest C., Funke R., Gage D., Galagan J., Gardyna S.,  
 RA Grant G., Hagos B., Headford A., Horton L., Howland J.C., Johnson R.,  
 RA Jones C., Kann L., Karatas A., Klein J., Lehotzky J., Lieu C., Locke K.,  
 RA Macdonald P., Margulis N., McEwan P., McGurk A., McKernan K., McLaughlin J.,  
 RA Meldrum J., Morrow J., Naylor J., Norman C.H., O'Connor T., O'Donnell P.,  
 RA Peterson K., Pollara V., Riley R., Roy A., Santos R., Severy P.,  
 RA Stange-Thomann N., Stojanovic N., Subramanian A., Talamas J., Tesfaye S.,  
 RA Tittell A., Vassiliev H., Vo A., Wheeler J., Wu X., Wyman D., Ye W.J.,  
 RA Zimmer A., Zody M.

Submitted (24-NOV-1999) to the EMBL/GenBank/DBJ databases.  
 Whitehead Institute/MIT Center for Genome Research, 320 Charles Street,  
 Cambridge, MA 02141, USA

On May 25, 2000 this sequence version replaced g1:7653382.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIPR

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: L4873

Center clone name: 30\_G\_23

Summary Statistics

Sequencing vector: M13; M77815; 95% of reads

4.86991327551701Chemistry: Dye-terminator Big dye; 100% of

Assembly program: Phrap; version 0.960731

Consensus quality: 152420 bases at least Q40

Consensus quality: 158679 bases at least Q30

CC	Consensus quality: 161216 bases at least Q20
CC	Insert size: 164000; agarose- <i>fp</i>
CC	Insert size: 163102; sum-of- <i>contigs</i>
CC	Quality coverage: 4.4 in Q20 bases; agarose- <i>fp</i>
CC	Quality cov.
CC	* NOTE: This is a 'working draft' sequence. It currently
CC	* consists of 17 <i>contigs</i> . The true order of the pieces
CC	* is not known and their order in this sequence record is
CC	* arbitrary. Gaps between the <i>contigs</i> are represented as
CC	* runs of N, but the exact sizes of the gaps are unknown.
CC	* This record will be updated with the finished sequence
CC	* as soon as it is available and the accession number will
CC	* be preserved.
CC	1 1226: <i>contig</i> of 1226 bp in length
CC	* 1227 1326: gap of 100 bp
CC	1327 2637: <i>contig</i> of 1311 bp in length
CC	* 2638 2737: gap of 100 bp
CC	2738 4238: <i>contig</i> of 1501 bp in length
CC	* 4239 4338: gap of 100 bp
CC	4339 6056: <i>contig</i> of 1718 bp in length
CC	* 6057 6156: gap of 100 bp
CC	6157 9089: <i>contig</i> of 2933 bp in length
CC	* 9090 9189: gap of 100 bp
CC	9190 12963: <i>contig</i> of 3774 bp in length
CC	* 12964 13063: gap of 100 bp
CC	13064 17269: <i>contig</i> of 4206 bp in length
CC	* 17270 17369: gap of 100 bp
CC	17370 25260: <i>contig</i> of 7891 bp in length
CC	* 25261 25360: gap of 100 bp.
CC	25361 32943: <i>contig</i> of 7583 bp in length
CC	* 32944 33043: gap of 100 bp
CC	33044 43011: <i>contig</i> of 9968 bp in length
CC	* 43012 43111: gap of 100 bp
CC	43112 55076: <i>contig</i> of 11965 bp in length
CC	* 55077 55176: gap of 100 bp
CC	55177 67728: <i>contig</i> of 12552 bp in length
CC	* 67729 67828: gap of 100 bp
CC	67829 83457: <i>contig</i> of 15629 bp in length
CC	* 83458 83557: gap of 100 bp
CC	83558 100009: <i>contig</i> of 16452 bp in length
CC	* 100010 100109: gap of 100 bp
CC	100110 116131: <i>contig</i> of 16022 bp in length
CC	* 116132 116231: gap of 100 bp
CC	116232 135453: <i>contig</i> of 19222 bp in length
CC	* 135454 135553: gap of 100 bp
CC	135554 164702: <i>contig</i> of 29149 bp in length.
XX	
FH	Key
FH	Location/Qualifiers
FH	
FT	source
FT	1. .164702
FT	/db_xref="taxon:9606"
FT	/organism="Homo sapiens"
FT	/clone="RP11-30G23"
FT	/clone_lbp="RPC1-11 Human Male BAC"
FT	1. .1226
FT	/note="assembly_fragment"
FT	1327. .2637
FT	/note="assembly_fragment"
FT	2738. .4238
FT	/note="assembly_fragment"
FT	4339. .6056
FT	/note="assembly_fragment"
FT	6157. .9089
FT	/note="assembly_fragment"
FT	9190. .12963
FT	/note="assembly_fragment"
FT	13064. .17269
FT	/note="assembly_fragment clone_end:sp6 vector_side:left"
FT	17370. .25260
FT	/note="assembly_fragment"
FT	25361. .32943
FT	/note="assembly_fragment"
FT	33044. .43011
FT	misc_feature

[illegible]

Yu, J. and Yang, H.  
Direct Submission  
Submitted (20-MAR-2001) Human Genomic Center, Institute of  
Genetics, Chinese Academy of Sciences, Datun Road, Beijing,  
100101, P.R. China  
On Mar 20, 2001 this sequence version replaced gi:8072552.

On Mar 20, 2001 this sequence version replaced gl:8072582.  
-----Genome Center

Center:Beijing Center  
Center code:Beijing  
Website:<http://hgc.igtp.ac.cn>  
<http://www.genomics.org.cn>  
Contact:hgc@igtp.ac.cn  
----- Project Information  
Center project name:18 project  
Center clone name: RP11-30G23

```
----- Summary Statistics -----
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; Er 55% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; Version 0.990329
Consensus quality: 499 bases at least Q40
Consensus quality: 605 bases at least Q30
Consensus quality: 674 bases at least Q20
Insert size: 692; sum-of-configs
Quality coverage: 2.80x in Q20 bases;sum-of-c-
```

Location/Qualifiers  
1. .166043

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"

```

```

/clone="RP11-30G23"
BASE COUNT 48570 a 35113 c 34446 g 47914 t
ORIGIN

```

7.28; Score 80.2; DB 9; Length 166043;

Best Local Similarity 73.0%; Pred. No. 5.4e-10;  
Matches 103; Conservative 0; Mismatches 38; Indels 0; Gaps 0.

Dy 137 agtgaactctgtgcccgaatgatgcaggagaagycatcccccaagcgtgtatactycaaga 196  
| | | | | | | | | | | | | | | | | | | | |  
Db 10231 AATAAACCTATTTTAAAAAATTGAGAGTAAGAACAATGCCAGGCATGTCCTCAGA 10290

[illegible]

```

QY      257  ggcagagttcacttcccaagg      277
          || ||||| ||||| |||
Db 10351  GGTAGAGCTCACTTCCCCAGG 10371

```

LOCUS	AC027126	185608 bp	DNA	PRI	06-MAR-2007
DEFINITION	Homo sapiens chromosome 3 clone RP11-586C12 map 3p, complete				

AC027126.4 GI:13236635  
HTG.

Homo sapiens  
Eukaryota; M

1 (bases 1 to 185608)  
Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J.,  
Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, L.

Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, X., Luo, C., Luo, T., Niu, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H.,

Wang, J., Wang, J., Wang, J., Wang, L., Wang, R., Wang, X.,  
Wang, X., Wang, X., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X.  
Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M.,  
Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,  
Yu, J., and Yang, H.

Chromosome 3p genomic sequence  
Unpublished

Direct Submission  
Submitted (28-MAR-2000) Human Genomic Center, Institute of  
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing

3 (bases 1 to 185608)  
Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J.,  
Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D.,

Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhu, B., Zhu, N., Yu, J., and Yang, H.  
Direct Submission  
Submitted (06-MAR-2001) Human Genomic Center, Institute of

**JOURNAL** Submitted (06-MAR-2001) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China  
**COMMENT** On Mar 6, 2001 this sequence version replaced gi:1101246.

On Mar 6, 2001 this sequence replaced gi:8101246.  
-----Genome Center

```
Center:Beijing Center
Center code:Beijing
Website:http://hgc.igbp.ac.cn
http://www.genomics.org.cn
Contact:hgc@igbp.ac.cn
----- Project Information
Center project name:18 project
Center clone name: RP11-586C12
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; ET 5% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 187858 bases at least Q40
Consensus quality: 189056 bases at least Q30
Consensus quality: 189280 bases at least Q20
Insert size: 165608; sum-of-contigs
Quality coverage: 10.96x in Q20 bases;sum-of-contigs
```

Location/Qualifiers

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/chromosome="3"
```

BASE COUNT	53641	a	40367	c	39222	g	52378	t
ORIGIN								

ch	7.2%;	Score 80.2;	DB 9;	length 185608;
1 Similarity	73.0%;	Pred. No. 5.6e-10;		
103; Conservative	0;	Mismatches 38;	Indels 0;	Gaps 0;

OY 137 agtgaactltgtgcgaatgcaaggagaagcattccccaagtgtatactacaaga 196  
| | | | | | | | | | | | | | | | | | | |  
DB 54490 AATTAAGCTATTTTAAAAATTCGAAGCTAAGAACATTCGCCAGGCATCTCCTCACA 54549



OY 728 atacagcaataaccctgaccgacgcaagatatagacagctagagagctgc 787  
 Db 148553 GTTTTATGGAATTTATTTCCGAATAATCCCTGTGGACATGCTTCAGTATGACGCTGC 148494  
 OY 788 agagcttctctgtataaacaagctgacctactctc 827  
 Db 148493 TTCAATTAGAGTTGAATAGAAACAACACTCACCGTATTTTC 148454

RESULT 14  
 AC092762 169580 bp DNA HTG 26-JUL-2001  
 LOCUS Pan troglodytes clone RP43-10804, WORKING DRAFT SEQUENCE, 9  
 DEFINITION unordered pieces.  
 AC092762  
 VERSION AC092762.1 GI:15022031  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE chimpanzee.  
 ORGANISM Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.  
 1 (bases 1 to 169580)  
 Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L.,  
 Granite, S., Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E.,  
 Lee-Lin, S.-Q., Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B.,  
 Masello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J.,  
 Pearson, R., Prasad, A., Shevchenko, Y., Snyder, B., Stantipop, S.,  
 Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsugen, C., Vogt, J.L.,  
 Walker, M.A., Wetherby, K.D., Zhang, L.-H. and Green, E.D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 169580)  
 Green, E.D.  
 Direct Submission  
 Submitted (26-JUL-2001) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gathersburg, MD 20877, USA  
 ----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: http://www.nisc.nih.gov  
 Contact: nisc-mouse@nigri.nih.gov  
 ----- Project Information  
 Center project name: ano  
 Center clone name: 108004  
 ----- Summary Statistics  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 164694 bases at least Q40  
 Consensus quality: 165758 bases at least Q30  
 Consensus quality: 166599 bases at least Q20  
 Insert size: 14100; agarose-fp  
 Insert size: 13200; pulse-field-gel  
 Insert size: 168780; sum-of-ctgigs  
 Quality coverage: 14.40x in Q20 bases; agarose-fp  
 Quality coverage: 15.40x in Q20 bases; pulse-field-gel  
 Quality coverage: 12.04x in Q20 bases; sum-of-ctgigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 9 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 2898: contig of 2898 bp in length  
 \* 2899 2898: gap of unknown length  
 \* 2899 9905: contig of 6907 bp in length  
 \* 9906 10005: gap of unknown length  
 \* 10006 16835: contig of 6830 bp in length

FEATURES  
 source  
 \* 16836 16935: gap of unknown length  
 \* 16936 24424: contig of 7489 bp in length  
 \* 24425 24524: gap of unknown length  
 \* 24525 38758: contig of 14234 bp in length  
 \* 38759 38858: gap of unknown length  
 \* 38859 53065: contig of 14207 bp in length  
 \* 53066 53166: gap of unknown length  
 \* 53166 74767: contig of 21601 bp in length  
 \* 74767 74866: gap of unknown length  
 \* 74867 96506: contig of 21640 bp in length  
 \* 96507 169580: gap of unknown length  
 \* 96607 169580: contig of 72974 bp in length.  
 location/Qualifiers  
 1. 169580  
 /organism="Pan troglodytes"  
 /db\_xref="taxon:9598"  
 /clone="RP43-10804"  
 /clone\_1bp="RP43"  
 1. 2898  
 /note="assembly\_fragment"  
 2899. 9905  
 /note="assembly\_fragment"  
 10006. 16835  
 /note="assembly\_fragment"  
 16936. 24424  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:left  
 24525. 38758  
 /note="assembly\_fragment"  
 clone\_end:T7  
 vector\_side:left  
 38859. 53065  
 /note="assembly\_fragment"  
 53166. 74766  
 /note="assembly\_fragment"  
 74867. 96506  
 /note="assembly\_fragment"  
 96607. 169580  
 /note="assembly\_fragment"  
 BASE COUNT 53358 a 33302 c 32407 g 49685 t 828 others  
 ORIGIN

Query Match 5.8%; Score 64.4; DB 2; Length 169580;  
 Best Local Similarity 48.8%; Pred. No. 7.8e-06;  
 Matches 204; Conservative 0; Mismatches 211; Indels 3; Gaps 1;

OY 407 ttccatataatcattcaatttcaagcgaatcgaatcctgacatgccaataaccaca 466  
 Db 158446 TTCCATCTGGAATTTCAATTTCTACTTATTAATTAAGATATTAAGTCATACACACA 158505  
 OY 467 tctcaatcttcgaagcaaatcgggtgttgaagaacctgaagaacctgaatgtggtt 526  
 Db 158506 TATCACATATATCACTTAAGAAATATTCAGCTTGGAATATACACACACTCTTTTATATA 158565  
 OY 527 tcaactatcgaagagcattccctccagaatgtggaatgttgaagaacctgaagagctg 586  
 Db 158566 ACAATTACATTTGAATAATTTCTCACTGACCTAGTAAGTCTTGGAACCTTGGA--AAATT 158622  
 OY 587 attgtctggaatctagaattatgaatgagcctgaccccttgatgaatgaatgaagaag 646  
 Db 158623 TAAGTTGGGTTAAATAAATTAAGTTAAGACATATACACAGTACTGCTAGTTTAAATACT 158682  
 OY 647 ttcaatttgaatataatcgaagcaaatcttccagctgcccgaatctgtccctgcgga 706  
 Db 158683 TGAGGCTTCAATTTGGAATATATATGATTAACAAATATTTCTTAAGCTCTGCTTCC 158742  
 OY 707 tgcgaattgcagctggtgtgataatcgaagcaataaccctgacgacactgcgcaagata 766  
 Db 158743 TTCCAAAGTTAATTTTCACTGACCTTACTGGAACCTTAATTAAGCAGTTTGCAGAAATAA 158802  
 OY 767 tagacagctagagagctgacgagcttctctgtataaacaagctgacctact 824

```

Db 158803 TTAGGACCTTTAAATTTAGAAACACTTTTATGATCACAATATGCTTACCTTCT 158860
|||||
RESULT 15
AF359380 1658 bp mRNA PRI 12-JUL-2001
LOCUS Homo sapiens LANO adaptor protein (LANO) mRNA, complete cds.
DEFINITION AF359380
ACCESSION AF359380.1 GI:14701833
VERSION AF359380.1
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1658)
AUTHORS Saito,H., Santoni,M.-J., Jaulin-Bastard,F., Marchetto,S.,
Isardon,D., Adelaide,J., Birnbaum,D. and Borg,J.-P.
TITLE Lano, a novel LAP protein directly connected to MAGUK proteins in
epithelial cells
JOURNAL J. Biol. Chem. (2001) In press
REFERENCE 2 (bases 1 to 1658)
AUTHORS Saito,H., Marchetto,S., Birnbaum,D. and Borg,J.-P.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2001) Molecular Oncology, U119 INSERM, 27,
Boulevard Lei Roure, Marseille 13009, France
FEATURES
source
1..1658
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1658
/gene="LANO"
1..1575
/gene="LANO"
/note="belongs to the LAP protein family with mammalian
Erbin, Scribble and Densin-180 but contains no PDZ domain;
interacts with MAGUKs (DLG) and LAP proteins (Erbin);
contains LRR motifs, a LAP specific domain (LAPSD), and a
C-terminal MAGUK PDZ domain binding site"
/codon_start=1
/product="LANO adaptor protein"
/db_xref="GI:14701834"
/db_xref="GI:14701834"
/translation="MPCITPLMRCNRHVESIDKRHCSLYVPEITRYANSLEPLID
ANOLRELPQFQVLVLRKLGSLDNETQRLPETANMQLVELDVSRNEIPEISIS
FCALQVADSGNPLVRLPESFELQNTCTSYNDISLOSIPENIMYINLASLEIRE
NLITVLPDSLTQLRLELDLGNNEIYNLDESIGALLHLKMDLQGNQSLSELPOEIGN
LKNLLCLDVSENRLERLPEEISGLTSLDLVISQNLLETIPDGIGKIKTSILKVDON
RLTQLEPAVGECSLPELVLTENOLLTPKISIGIKKLSNLMNDRKIVSLPRKIGSC
CSITVECVRNRLRTRIPAVSQATELHVLVDVAGKRLHLPLSLTALKLKLWLSNDNS
QPLITFQITDVTYTGKILTCVLPLPQSEPTCOENLPCGALENLVNDVSDENMR
AVNRVSAIRFVEDEKEDENETRTLRLRAPHPGELHMKRYENLRNDMNAAGLDS
NKNEVNHAIIDRVTTSTV"
BASE COUNT 483 a 383 c 378 g 414 t
ORIGIN
Query Match 5.7%; Score 63.6; DB 9; Length 1658;
Best Local Similarity 46.4%; Pred. No. 5.7e-06;
Matches 244; Conservative 0; Mismatches 279; Indels 3; Gaps 1;

```

```

QY 563 attgtaaatcttagagagactggaattgtctggaatcttagaattatgagctgcct 622
Db 564 CCCTCTTACATCTTAAAGATCTCTGTGGATGGAATCAACTGTCA--GATTAACCTC 640
QY 623 ttgaattagtaattggaagcaagttacattgtatagatattccagcaacaagttcca 682
Db 641 AGGAATATGGAATCTGAAAGAACCTGCTGTGTTTATGATGTCTCTTAAACAGGTGGAA 700
QY 683 gtgcccacatctgtgctctgaggaatgtgcaattgtcagtggtgatactcagcaacaata 742
Db 701 GACTTCCTGGAAGAAATCAGTGGCGCTGACTTATTAACGGATTACTCAATTCGCCAGAACT 760
QY 743 acctgacccgacctgcgcgaagatatagacaggtcagagagctcagagcttctctgt 802
Db 761 TATTGAAACGATTCCTCGATGGCATTGGAACCTTAAAGAACTGTCAATCTTGAAGGTGG 820
QY 803 ataaacaagtgtagcctacccttccattccattccattccattccattccattccattcc 862
Db 821 ATCAGAAATAGACTCACACAGTTGCTTGAAGCAGTTGGGGAATGTGAAAGTCTCACTGAGT 880
QY 863 tagtcgtagtgaggagaccatttggtgagagctcccaactgccttg 908
Db 881 TAGTCTTACAGAAATCAGTCCCTGACCTGACCTGCTTAAAGCATTTGG 926

```

Search completed: February 26, 2002, 10:52:07  
Job time: 19404 sec